Methodology to associate seismic data between Petrobras' Proprietary DB and Open-Works®' projects

Ricardo Álvares dos Santos (Halliburton), Renan de Jesus Melo (Petrobras), Filipe dos Santos Sá (Halliburton) , Bruno Auad David (Halliburton), Renan Jandre Esteves (Halliburton), Leandro Pacheco Vaz (Halliburton)

This paper was prepared for presentation during the 14th International Congress of the Brazilian Geophysical Society held in Rio de Janeiro, Brazil, August 3-6, 2015.

Abstract

This work shows the methodology we developed to compare seismic information stored in two different digital formats: SEG-Y (.sgy) and Compressed Landmark® format (.cmp).

In addition, we used DecisionSpace Integration Server (DSIS) to create reports comparing two different databases with seismic data in each format.

Introduction

Petrobras' seismic data management has two bigger processes: data stock at Petrobras' Proprietary DB (PPDB) - accessed by a house developed web browser application - and the loading of this data into interpretation platforms (mainly Landmark OpenWorks - OW).

The loading stage transforms a seismic data from SEG-Y Exchange digital format (Rev 0 or Rev1) into a proprietary (compressed or not) format. In the house, we use Landmark's CMP compressed format with a fidelity rate of 99.9%. When a geologist or geophysicist needs a seismogram, three steps had to be taken: 1) send a seismic volume demand for responsible area; 2) Download SEG-Y format of seismic data from PPDB; 3) To load this data into OW, for example.

This workflow starts on demand (step 1). If a request is not made, a specific seismic volume will not be loaded. Thus, PPDB and OW's database are not synchronized.

In the past, we tried manually make a report that would have shown which data was missing and which one was loaded in both places. Although, PPDB has a big amount of data. Therefore, we needed a software solution that can make this work faster.

We choose Landmark DecisionSpace Integration Server (DSIS) to create an environment for data comparison and DecisionSpace Web Framework (DSWF) to create a web service where our collaborators in the department could visualize the comparison reports.

Method

Every seismic volume at Petrobras's Proprietary Database is stored in SEG-Y Petrobras' extension format and each one has (at least) three metadata files: PROC, GRID, AMPL_SCAN.

• PROC is an ASCII copy of EBCDIC header

• GRID is a text file with grid and cartographic information

• AMPL_SCAN is a text file with volume's amplitude information

These seismic volumes are not compressed. Generally, one single volume occupies 100 GB to 600 GB. When one of these files is loaded into OW, it's compressed using a fidelity rate of 99.9%, which reduces its size to 10 to 50 GB and changes its digital format from SEG-Y to Landmark's CMP format.

This process has two important points: 1) less space is needed for the same seismic information; 2) compression could change seismic amplitudes registered in the original format. Because of this change in amplitude, the comparison of these two files is not so easy.

We noticed that when we load a SEG-Y in an OW's survey with the same grid parameters and use a fidelity rate of 99.9%, the resulting volume has amplitudes a little bit different due to compression process but volumes' maximum and minimum amplitudes remain the same.

At PPDB, we observe seismic volumes from different areas at the country, with different seismic processing techniques and parameterization. Unless seismic amplitudes were trimmed during the processing stage, it is high unlikely to find two seismic volumes with exact the same maximum and minimum amplitude values. This result is valid for both PPDB and for the OW.

Therefore, we used these two criteria to compare data between PPDB and OW: grid information and extreme values. If they match, then we associate a correspondence between them.

PPDB splits seismic information in two parts: acquisition metadata stays in an Oracle instance and seismic data (SEG-Y, PROC, GRID, AMPL_SCAN, etc.) is stored on a file system linked to the oracle database. PPDB web app access these information (Figure 1).

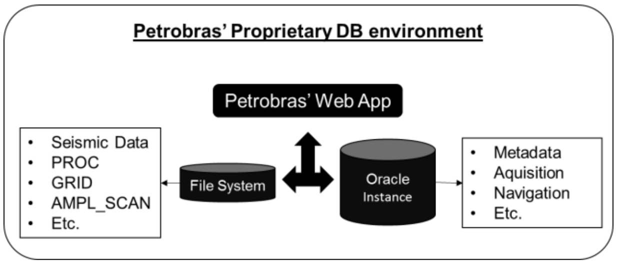

Figure 1 - Petrobras' Proprietary DB

DecisionSpace Data Server (DSDS) provided the technology for data federation between databases. Once aimed metadata files were on a file system, we needed to download PROC, GRID and AMPL_SCAN and populate a virtual database (VDB).

We divided our methodology in 5 steps of implementation:

1) IT infrastructure: server configuration and software installation

2) Download from PPDB: PROC, GRID and AMPL_SCAN

3) Loading PPDB's data into a virtual database (VDB)

4) DecisionSpace Data Server connection between VDB and OpenWorks database

5) DecisionSpace Web Framework: comparison results are shown in a webpage.

Steps 1 and 2 occurred at the same time in different machines. Figure 2 illustrates implementation process as a hole.

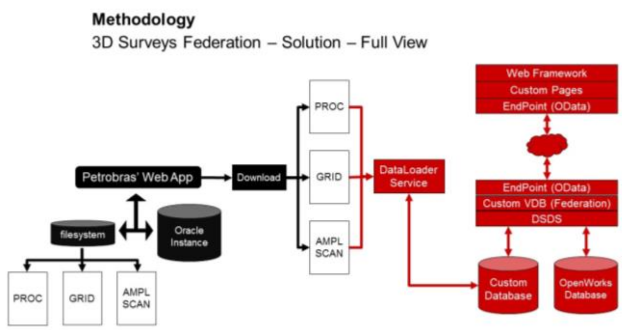

Figure 2 - Comparison: IT implementation

Webpage navigation

In OW, surveys and volumes are organized by districts and DecisionSpace Web Framework provided a web environment to access comparison results. Bellow we show the webpage used: A) choose a district; B) graphs with surveys, grids and quantitative volumes information; C) Volumes comparison between PPDB's data and OW's surveys with grid match.

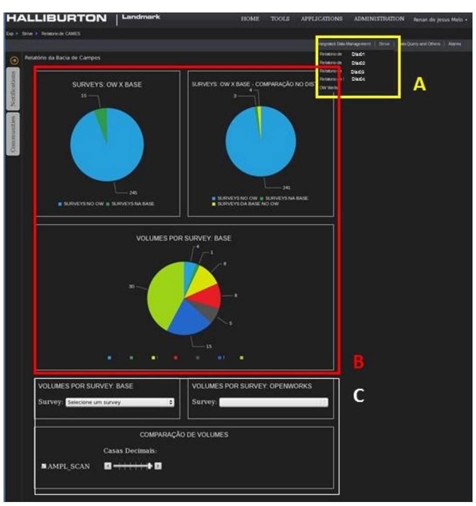

Figure 3 – Comparison: result's webpage

Webpage part C brings a table that shows volumes matches between PPDB's grid and OW's surveys. In addition, an AMPL_SCAN bar controls numerical precision in the comparison (Figure 4)

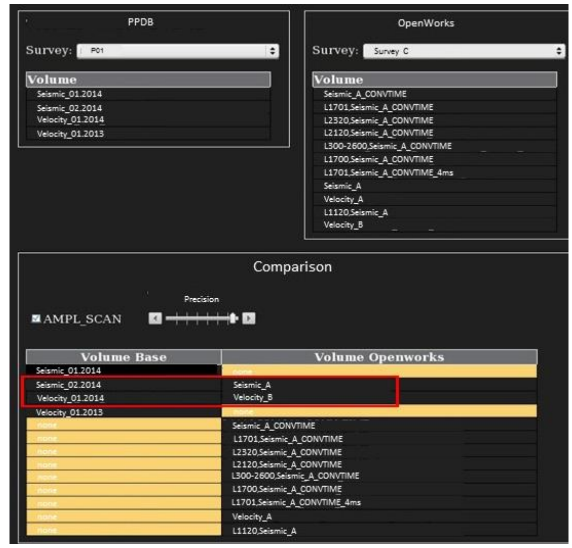

Figure 4 – Webpage part C: numerical precision control and comparison results

Results

Petrobras's official repository has thousands of seismic processing programs and, generally, each one has several seismic volumes. Because of this huge number of seismic files, we chose only seven seismic processing programs as a test.

Although seven processing programs is a minimum part of PPDB, we took care to choose seismic programs from different Brazil's sedimentary basins with different year of acquisition, sizes and processing techniques.

Table 1 - Comparison dataset: general information

Basin	Processing Number	PPDB's volumes	OW's site volumes
Almada	01	04	973
Campos	02	01	2786
Campos	03	08	2786
Ceará	04	08	1069
Almada	05	05	973
Santos	06	15	3439
Santos	07	29	3439

Observing Table 1 we can have a grasp of this herculean job: as an example, find one of processing #7's 29 volumes among 3439 OW's volumes. Usually, OW's database has more seismic volumes than PPDB because of volume attributes generated by the geophysicists. Then, reverse seeking consulting which volumes generated by the interpreter have not been stored in the database is also an extremely difficult task.

Our first match difficulty was that PPDB organizes volumes with different grids inside the same processing number, but OW's does not allow this organization in the same seismic survey. Figure 5 shows the grids founded in each processing number.

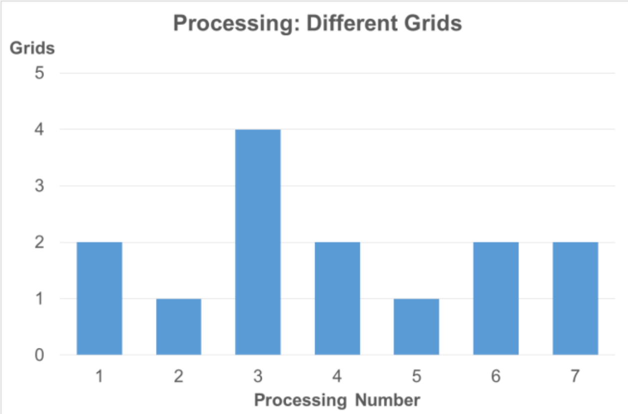

Figure 5 - Number of grids for each processing program

We compared grid by grid, volume by volume, using the methodology described above. Our results are shown in the table below:

Table 2 – General results for dataset comparison

PPDB			COMPARISON	OPENWORKS		RESULTS	
Processing	Grid	Volumes		Survey	Volumes	Volumes Founded	Volumes Not Founded
01	1	1		none		0	0%
01	2	3		Survey A	4	0	0%
02	1	1		Survey B	67	0	0%
03	1	5		Survey B	67	3	60%
03	2	1		none		0	0%
03	3	1		none		0	0%
03	4	1		none		0	0%
04	1	4		none		0	0%
04	2	4		Survey C	11	2	50%
04	2	4		Survey D	2	0	0%
05	1	5		Survey E	8	3	60%
06	1	12		Survey F	25	7	58%
06	1	12		Survey G	6	3	25%
06	2	3		Survey H	9	0	0%
07	1	8		none		0	0%
07	2	21		Survey I	66	0	0%
07	2	21		Survey J	0	0	0%
07	2	21		Survey K	31	17	81%

(red)	Grids not founded in OW
(yellow)	Grids founded in OW, but volumes not loaded
(green)	Grids and volumes loaded

Observing Table 2 we noticed three general groups and they were differenced by a color code:

• Red: no match were found between program and OW's survey

• Yellow: grid match but no amplitude similarity

• Green: grid and amplitudes matched. Last column shows match rate in percentage, i.e., number of matches / number of seismic volumes.

From PPDB, we used 70 seismic volumes. Quickly, the methodology we developed showed that 35 seismic were loaded (50%). Volumes that did not match with any of OW's data, three cases are possible: 1) they were not loaded, 2) they were loaded inside the wrong grid or 3) they were loaded without a 99.9% fidelity rate.

Another parameter that may influence the amplitude comparison is data precision or data quantization. Example: seismic amplitudes were acquired at 32 bits but loaded into OW at 8 bit. Although this possibility exists, we didn't check it.

Conclusions

In this production environment, the comparison methodology has proved to be a fast way to compare seismic data in different digital formats because it uses only metadata instead of seismic volume itself.

The ideal situation to test our methodology would be in a controlled DB were every volume is loaded in a specific grid. Thus, the low match rate is acceptable as we apply the methodology in a production environment whose loading rules changed all over the years.

DecisionSpace Integration Server provided a fast and web-based solution to visualize comparison results in our methodology. Almost the same methodology can be applied to 2D seismic data and, further criteria could be extend to well data comparison.

Acknowledgments

We thank the IT teams of Halliburton and Petrobras that were dedicated to the installation of DecisionSpace Integration Server (DSIS).

References

We didn't find a similar paper in the subject.